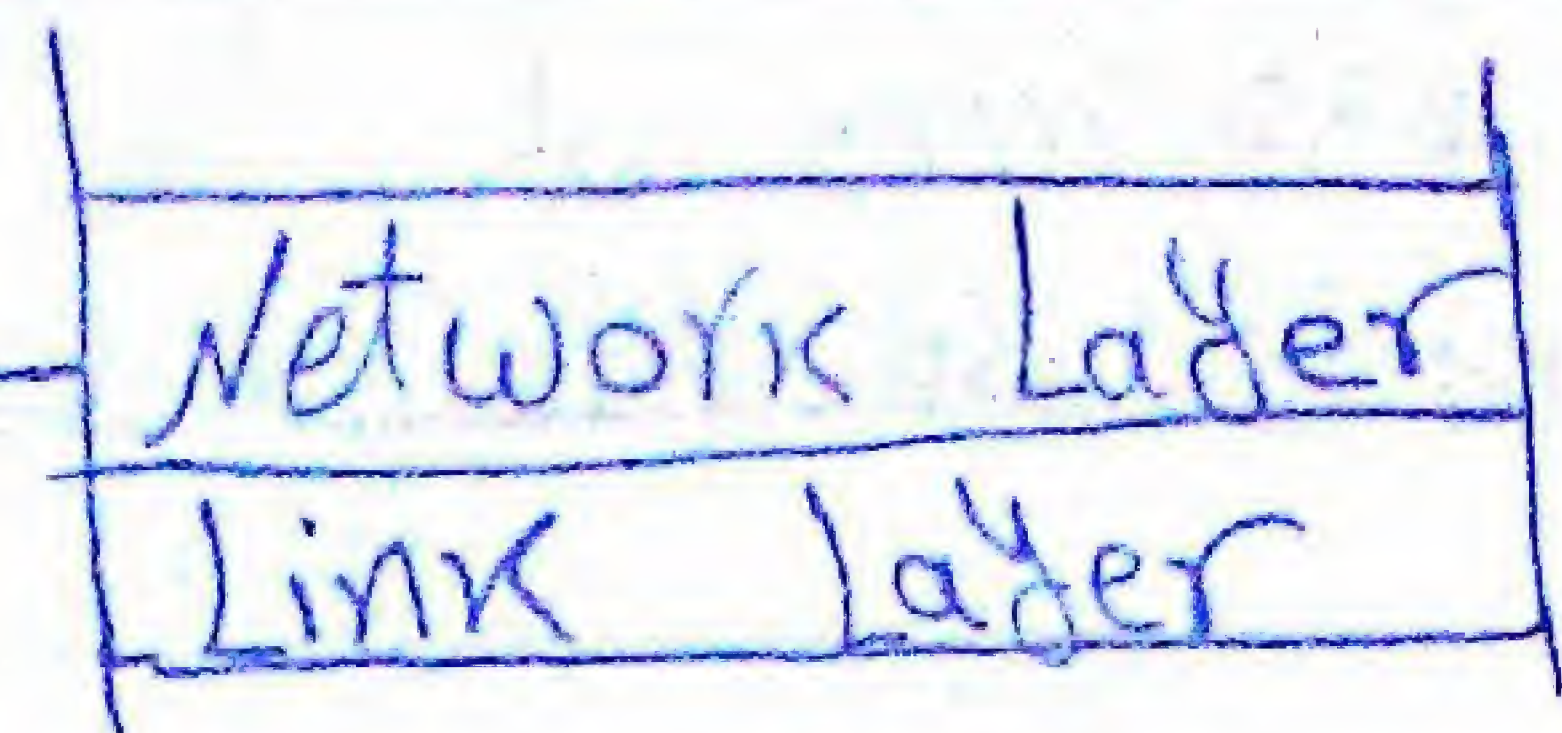


## ch5 - Link layer

2 hosts

Communication بين



Protocol stack

### \* Types of link-layer channels:-

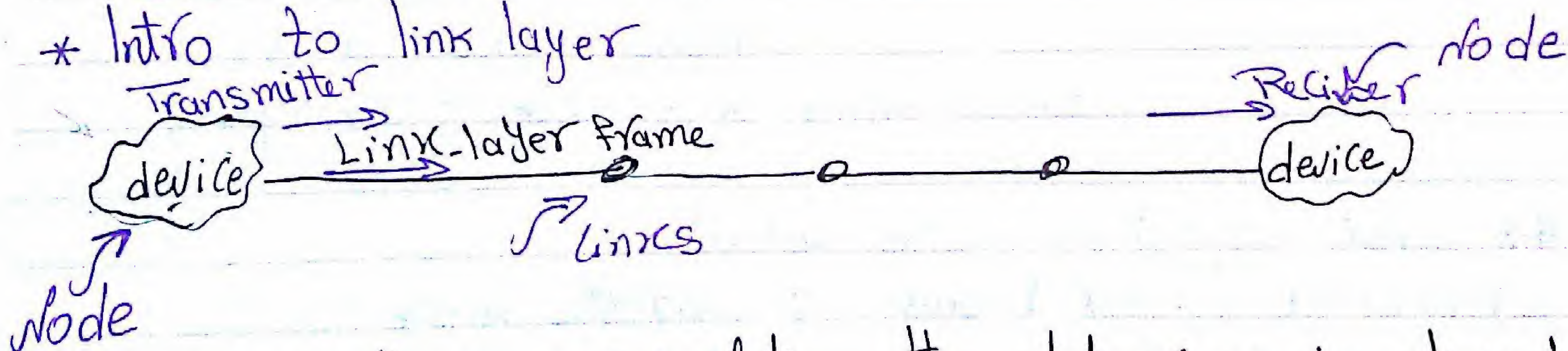
#### 1- Broadcast channels

توصيل بين multiple hosts خلال Satellite, wireless LANs, HFC  
وبما أنهم multi hosts فحتاج medium access protocol لنظم transmissions  
بينهم وعند طاليتين ممكن يكون في Central Controller للتنظيم لاول  
hosts بنفسها تنظم الـ trans. بينهم.

#### 2- Point-to-Point links

Ex: Router to Router - PC to Ethernet switch.  
وبما أنهم PPP فتظيم الـ access والنظم يكون بسيط

### \* Intro to link layer



\* Transmitter encapsulates the datagram in a link layer frame and transmits the frame into the link.



## \* Services of the link layer:

\* الوظيفة الأساسية لها أن تنقل الـ datagram من src. node إلى des. node

### 1\* Framing

مسئولة أن تقبل وتنقل الـ datagram إلى الـ link قبل أن يذهب إلى الـ link-layer frame . Encapsulation

Frame : 

Data field	header fields
------------	---------------

### 2\* Link access

لتنسيق الوصول إلى الـ medium access control protocol (MAC) يحدد أنواع

ويكرف عن عملية نقل الـ frame إلى الـ link

- ✓ Point-to-Point links → MAC's simple or don't exist.
- ✓ Broadcast link → (multiple access problem), MAC Protocol's needed to coordinate frame transmission.

### 3\* Reliable delivery

من خلال acknowledgements & retransmissions

لتنسيق جميع الـ links التي فيها معدلات عالية من الأخطاء errors

زى الـ wireless links

و يبقى الاتصال ليها حالة الـ wired links

### 4\* Error detection and correction

- bit errors <sup>may</sup> occur because of magnetic noise

- No need to forward datagram that has error bits  
So link layer protocols provide mechanism for error detection and correction.

- By, transmitting node include error detection bits in the frame, Receiving node check these bits.

- Error detection in the link layer's implemented by hardware.



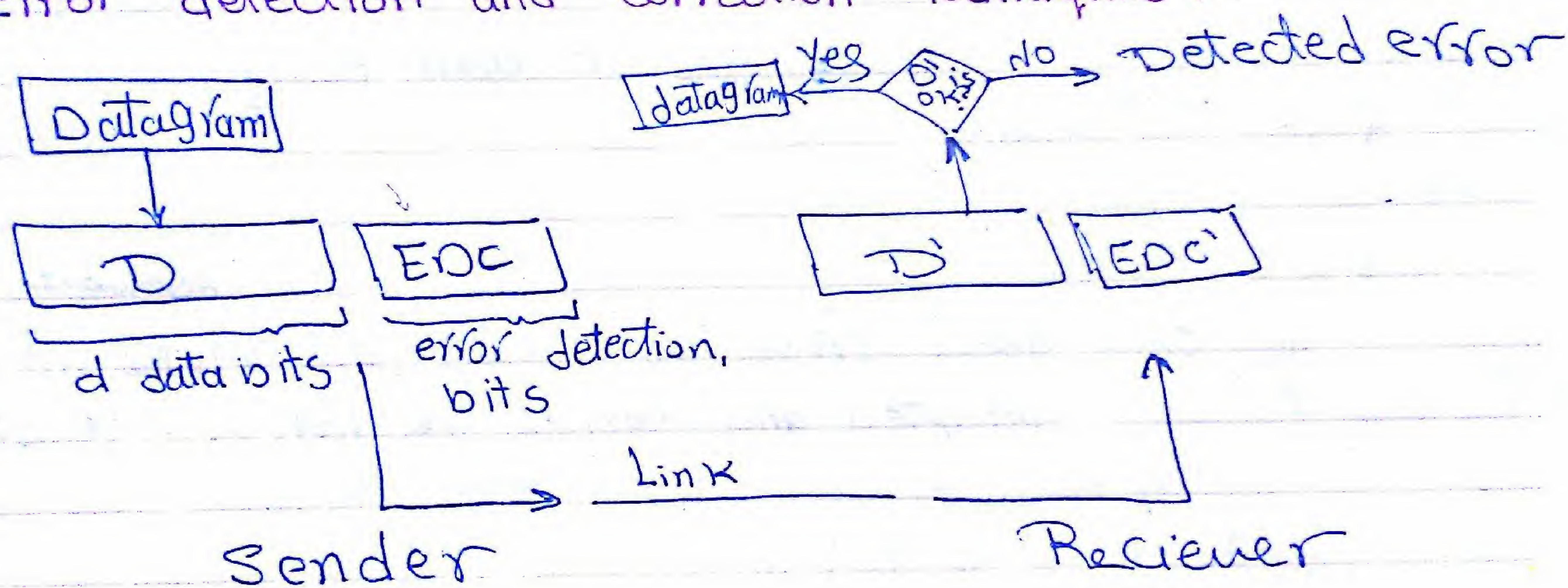
\* Where's the link layer implemented?

- Link layer's implemented in a network adapter (Network Interface Card (NIC)).

- link layer controller: ~~the~~ chip, It's part of network adapter, implements link layer services.

- The link layer's combination of HW and SW.

\* Error detection and correction techniques:-



- Even with the use of error detection bits there still may be undetected bit errors

→ Error detection techniques:

- ↳ Checksum methods (more used in transport layer)
- ↳ Cyclic redundancy checks ( ~ ~ ~ Link layer)
- ↳ Parity checks



\* Parity checks

↳ Single Parity bit for error detection.

- In even parity schema: no. of 1's in  $(d+1)$  bits is even
- In odd " " " " " " is odd

\* ReGener only count \* of 1's in the recieved  $(d+1)$  bits

\* في ال Even Parity schema : لو عدد الوايت فردى ← يضاف الاقل في bit \* حصل فيها error أو عدد فردى من ال bits حصل لهم error .

\* ولو كان عدد زوجى من ال bits حصل لهم error ← undetected error

0 1 1 0 1 1 1  
 └──────────┘     ↓  
 d bits           bit parity

→ one-bit even Parity

↳ 2 d Parity

determine

Not only detect error has occurred, but can the bit that was corrupted and correct the error. (Single bit error)

2d even Parity

\* 2d Parity Can detect (but not correct) any combination of 2 errors in a packet.

\* The ability of the receiver to both detect and correct errors is forward error correction (FEC)



### \* Check Summing

→ at sender

يجمع ال Segment bits وبتا 16 bit

ويجمع ال Segment bits وبتا 1's Complement  
وتخزنها في ال checksum bit

→ at receiver

يجمع ال Segment bits ال وبتا  
ويجمع ال checksum bit وبتا 1's Complement  
لو الناتج طلع 0 فكل شيء سليم  
لو الناتج وبتا 1 فكل شيء غير سليم

→ why's checksum used at transport layer and cyclic redundancy check used at link layer?

• Transport layer's implemented in SW so it's important to have a simple and fast error detection schema such as checksumming

• Link layer's implemented in HW in adapters, that can perform more complex CRC operations.

### \* Cyclic Redundancy Check (CRC)

- CRC → Polynomial Codes

$\underbrace{\quad\quad\quad}_{d\text{-bits}} \quad \underbrace{\quad\quad\quad}_{r\text{ bits}}$   
D: Data to be sent      R: CRC bits      (Bit Pattern)

$r+1$  bit pattern → Generator G (Leftmost (Significant) bit's 1)



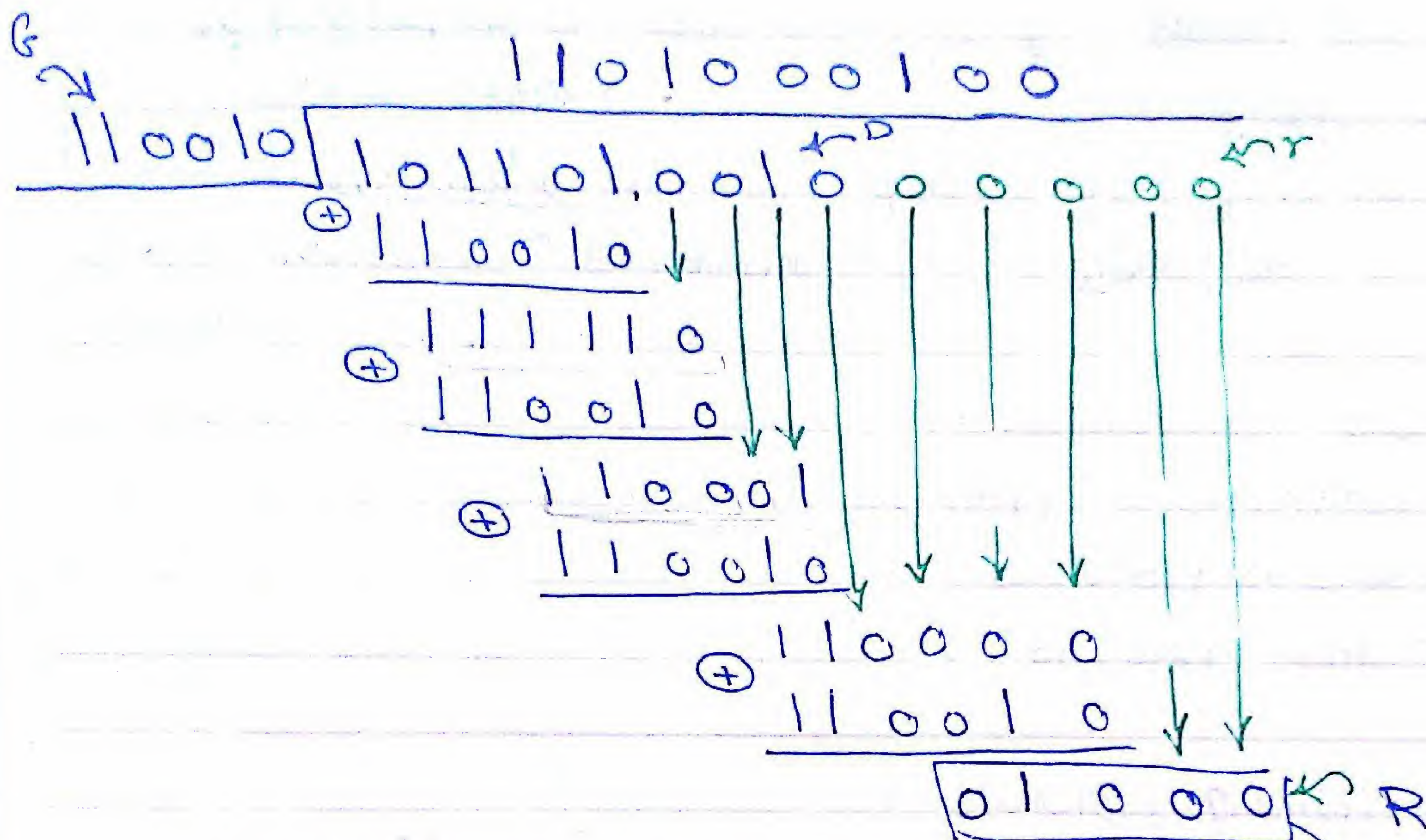


Example

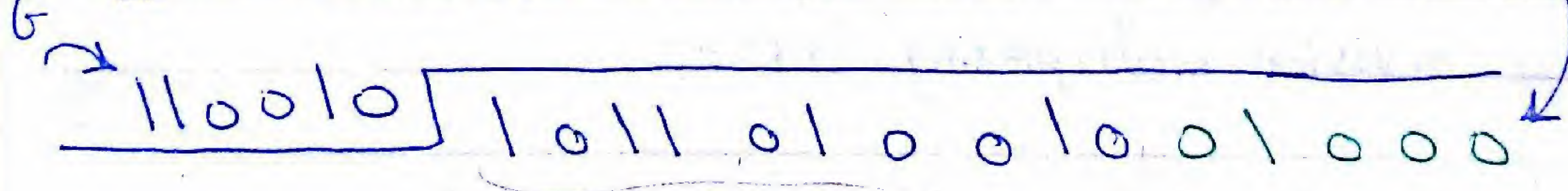
$D = 1011010010$  ,  $d = 10$  bits

$G = 110010$  ,  $r = 5$  bits

at sender



at Receiver



Reminder  $\rightarrow$  if  $= 000 \rightarrow$  No errors happen





## \* Multiple access links and protocols

\* P2P link protocols: PPP - high level data link control (HDLC)

\* broadcast link technologies: Ethernet - wireless LANs

← لو في أكثر من 2 node حاولوا يبعثوا frames في نفس الوقت في broadcast link  
 باقى ال nodes مستقبل في نفس الوقت وهنا يحصل Collision و frames  
 هتكون lost و Bw هتكون wasted وبالتالي فينا جيب ترتيب  
 Transmission Coordination و دور ال MAC protocols

## → MAC protocols categories:-

↳ channel Partitioning protocols

↳ Random access protocols

↳ Taking turns protocols

## \* channel partitioning protocols

→ used techniques to partition a broadcast channel's BW:-

① ↳ Time-division multiplex (TDM) ② ↳ Code division multiple access (CDMA)

③ ↳ Frequency-division multiplexing (FDM)

\* TDM:

- يقسم ال time إلى مجموعة Time frames و Time frames إلى

$N$  time slots ، كل time slot بيحصل assign لواحدة من ال

$N$  nodes ، ليراعوا و هما ليتنا رخوا ال slot size (التي هتكون

كافية لنقل Packet

Time  
↓  
frames  
↓  
N-slots

- Adv: TDM eliminates collisions , and fair  
 each node take  $R/N$  bps each time frame

- Disadv: \* ~~over~~ a node is limited to an avg rate of  
 $R/N$  even when it's the only node with packets  
 to send.

\* A node must always wait for its turn in trans. seq.



### \* FDM

يقسم ال channel لfrequencies كثافة كل واحدة لها  $BW \leftarrow R/N$  ويصل  
 assign لfreq. لو احدى ال  $N$  nodes  
 كأنه يصل  $N$  channels بعدل  $R/N$  bps

له نفس مزايا وعيوب ال TDMA

### \* CDMA

يصل assign لdifferent codes لكل node، ال node يوصل encoding للبيانات  
 للرسالة بال code وكل receiver يتقبل sender's encoded data bits (ال receiver  
 عنده عارف ال code ال Sender)

يستخدم في ال military Purpose

### \* Random access Protocols

Transmitting node always transmits at the full rate of the channel,  
 $R$  bps. When there's a collision, each node retransmits  
 its frame after ~~different~~<sup>random</sup> delays

→ Used Random access protocols

① → ALOHA Protocols

② → CSMA Protocols

③ → Ethernet Protocol

### \* Slotted ALOHA

assume that frame  $\rightarrow L$  bits, time slot size  $\rightarrow L/R$ ,  
 transmission  $\rightarrow$  beginning of slot, nodes're synchronized,  
 if frames collide, all nodes detect collision before end of  
 slot.

ال node يثبت ال frame في بداية ال time slot، إذا ~~تصادم~~ حصل  
 collision يبعث ال frame ويبدأ ~~بجهد~~ ال frame الى بعد  
 إذا ~~تصادم~~ حصل detect collision تبدأ ال frame



مرة ثانية خلال subsequent slot بـ Probability  $P$

- Adv:

- \* allows a node to transmit continuously at the full rate  $R$  when that node is the only active node.
- \* highly decentralized, as each node detects collisions and independently decides when to retransmit.
- \* Slotted ALOHA works well when there's only one active node.

Figure 5.10 → 451

Notes on Figure:

- Successful slot's in which exactly one node transmits.
- if each node were immediately retransmit after each collision, → efficiency = 0

- max. efficiency of slotted ALOHA:-

assume:

Probability  $P$  Collision بـ node في وقت إرساله

$N$  nodes →

احتمالية أن slot تكون successful slot هي احتمالية أن node

واحدة تعطي transmit والباق  $N-1$  nodes لا تعطي

$$P_{\text{success slot}} = P \times (1-P)^{N-1}$$

∴ There're  $N$  nodes

$$P = NP(1-P)^{N-1}$$

$$\therefore \text{Efficiency} = NP(1-P)^{N-1}$$

$$\therefore \text{max efficiency} = 1/e = .37$$

$$\therefore \text{effective transmission rate} = .37R \text{ bps}$$



### \* ALOHA

- First ALOHA, decentralized, unslotted.

- node يوصلها ال Frame بتة transmit فوراً ؛ لو ال Frame حصل ال Collision ← ال node هتة Retransmit فوراً باحتمالية  $P$  خلاف كدة ؛ ال node بتتقصر ال Frame trans time وبتكون كدة ثبتت ال Frame باحتمالية  $P$

Figure 5-11 → 453

- max efficiency of Pure ALOHA :-

$$P = P(1-P)^{2N-1} = P(1-P)^N (1-P)^{N-1}$$

$$\therefore \text{max efficiency} = 1/2e = \frac{1}{2} \text{ of slotted ALOHA.}$$

### \* Carrier sense multiple access (CSMA)

Rules - node listens to channel before transmitting → Carrier sensing  
 - a transmitting node listens to the channel while it transmitting  
 if it detects that another node is transmitting an interfering frame, it stops transmitting and waits a random amount of time before repeating the sense and transmit when idle cycle → Collision detection

These rules 're in CSMA and CSMA/CD

→ Why collisions still occur?

Figure 5-12

### \* channel propagation delay

الوقت اليتا خده ال signal عشان يوصل ال Propagation من node لاخرى  
 وكل ما ال delay بيزيد ، فرصة ال node انزل تغل carrier sensing ل node ثانية بديان بتقل



## \* CSMA / CD

\* ~~nodes~~ when nodes performs collision detection, it aborts their transmission a short time after detecting a collision, to abort them not transmitting a useless frame.

\* Operation :- Figure 5.13 → 456

- 1- Prepare datagram from NL.
- 2- if the adapter (in node) senses the channel's idle, it transmits the frame, if do, channel's busy, it ~~wait~~ waits until it senses no signal.
- 3- While transmitting, the adapter monitors the channel for signal energy coming from other adapter.
- 4- if detect energy, it aborts its transmission
- 5- After aborting, it waits random amount of time and then returns to 2

→ Why we need a random amount of time (Not fixed)?

\* Collision وبقدره  
انتظروا نفس المدة هيصل انهم Colliding forever

\* Random amount of time → Back off time

→ what's a good <sup>time</sup> interval to choose Back off time?

\* Large time + small \* of colliding nodes → They wait a large time.

\* small time + large \* of colliding nodes → nodes'll again collide.

\* Preferable: → short time ← \* colliding nodes're small  
→ long time ← \* colliding nodes're large



## \* Binary exponential backoff algorithm

↳ used in Ethernet and DOCSIS cable network to solve this problem.

↳ When transmitting a frame that has already experienced  $n$  collisions, a node chooses the value of  $K$  at random from  $\{0, 1, 2, \dots, 2^n - 1\}$ .

↳ in Ethernet, actual amount of back off time =  $K \cdot 512$

↳ The size of the sets from which  $K$  is chosen grows exponentially with # of collisions. So it's called exp. backoff algo.

## \* Taking turns protocols

→ Examples of used protocols:

① ↳ Polling Protocol

② ↳ Token passing Protocol

### \* Polling Protocol

\* يتطلب أنه يكون واحدة من nodes هو الـ master

\* الـ master يت Poll كل node ع هيئة Round Robin Fashion

بعض أوقات لا node الأول message بأفائة، تعجز transmit  
 frames max معين و لا الأول تخص تبعث الثانية وهكذا  
 والـ master ليف، يعرف الـ node خلصت ولا لا بالتتابع  
 channel

### \* Adv:

- Eliminate collisions and empty slots
- higher efficiency.

### \* Dis adv

\* Protocol introduces a Polling delay (الوقت التي بضع فيها الـ node الـ Poll الـ nodes <sup>②</sup> inactive) (ببعض الـ nodes الـ Poll الـ nodes <sup>②</sup> inactive)



\* If the master node fails, the entire channel becomes inoperative.

\* Examples of Polling protocols :-

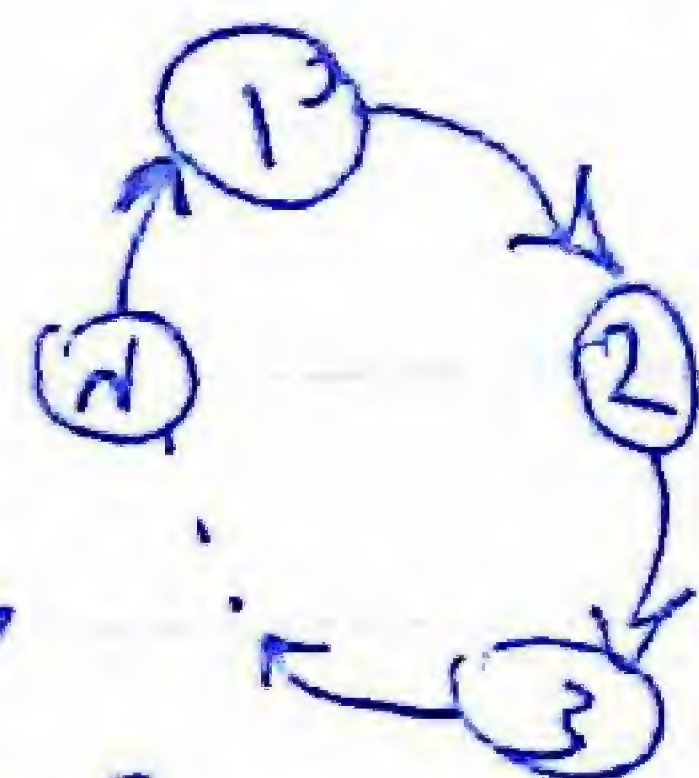
↳ 802.15 Protocol

↳ Bluetooth Protocol

\* Token Passing Protocol

\* No master node

\* Special Purpose frame known as a Token is exchanged among the nodes in some fixed order.



\* node التي يملك token يستطيع إرسال frames أو transmission أخرى غير frame token إلى node التالي.

→ Adv:

\* Token passing is decentralized and highly ~~effective~~ efficient.

→ Disadv:

\* The ~~set~~ failure of one node can crash the entire channel.

\* If a node accidentally neglects to release the token, then some Recovery procedure must be invoked to get the token back in circulation.

Examples of token Passing Protocol:

↳ Fiber distributed data interface (FDDI)

↳ IEEE 802.5 token ring Protocol



# MAC Protocols (summary)

Channel Partitioning

- TDM
- FDM
- CDMA

Random access

- slotted CSMA
- ALOHA
- CSMA/CD

Taking turns

- Polling
- Token Passing





## \* Cable access network (case study)

~~Fig 5.14 P 461~~

Fig P 38 slides

\* Connect Cable modems to Cable modem termination system (CMTS) at the cable network headend.

## \* Data over Cable Service Interface Specification (DOCSIS)

↳ specifies cable network architecture and its protocols.

↳ uses FDM <sup>to divide</sup> upstream & downstream into multiple freq.

↳ In downstream: CMTS → all Cable modems &

There is no multiple access Problem.

(CMTS → modems)

↳ In upstream: (challenging) < collisions can occur.

(modems → CMTS)

↳ upstream uses TDM to divide into intervals & time, each containing a sequence of mini-slots & during which modems can transmit to CMTS.

↳ CMTS give permission to cable modem to transmit by sending control message (MAP message)

↳ How CMTS know which cable have data?!

By having cable modems send mini-slots request frames to the CMTS during a special set of interval ~~time~~ mini slots that're dedicated for this purpose. They're transmitted in random and collision can occur btw them.

جيب دلوقتى ال cable modem من عارف؟ ذاك ال link هو اللي busy وة

ال frames حاصها Collision

فهو هيفرض انت لو مستقبل ردت من ال CMTS لبتاكد حصل Collision



في حالة انه حصل Collision في ال Cable modem  
 عن ان يعيد ارسال ال frame  
 Back off delay  
 ليستون

\* Switched Local area networks

→ Link layer addressing and ARP





## Switched local area networks

### \*Link layer addressing and ARP

Hosts and routers have link layer addresses, network layer addresses as well.

Address resolution Protocol (ARP) provides mechanism to translate IP addresses to link layer addresses.

### → MAC addresses

Hosts and routers' adapters have link layer addresses (adapters → network interfaces).

link layer address of interface یعنی آن کی

link layer switches' interfaces don't have link layer addresses.

link layer address's called LAN address, physical address, MAC address.

MAC address is 6 bytes long (giving  $2^{48}$  possible address), expressed in hexadecimal notation.

MAC address is unique for each adapter. IEEE manages the MAC address space.

Adapter's MAC address has a flat structure and doesn't change no matter the adapter goes.





### - Scenario (Sending & receiving Frames)

When an adapter wants to send a frame to some destination adapter, it inserts destination MAC address into the frame and then sends the frame into LAN.

A switch broadcasts an incoming frame onto all of its interfaces. So an adapter may receive a frame that isn't addressed to it.

When an adapter receives a frame, it'll check to see whether the des. MAC in the frame matches its own MAC address. If it matches, the adapter extracts the encapsulated datagram and passes the datagram up the protocol stack. If it doesn't match the adapter discards the frame.

If a sending adapter doesn't want all the other adapters to receive the frame, the sending adapter inserts a special MAC broadcast address into the destination address field of the frame.

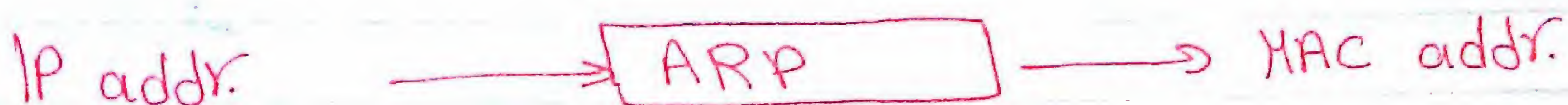




\* what ARP does?

Translate between link layer address and network layer address.

لوف host A عزيفت Frame ل host B لاطومية ال IP ل B  
ففيضيف ل Frame ال destination MAC addr فيضصل عليه  
عن طريق ال ARP



- ~~req~~ - ARP resolves IP addresses only for hosts and router interfaces on the same subnet.

### \* How ARP works?

- Each host and router has ARP table.

## ARP table

IP addr. | MAC addr. | T.T.L  
(when each mapping will be deleted)

→ (sending & Receiving at the same subnet)

Scenario → (IP → MAC mapping not found in the table)

\* لو ال table مضموش mapping ل IP معين ، ال Sender بيرسل ARP Packet كس ال nodes الموجودة في ال Subnet ، وكن node هتقارن ما بين dest. IP addr. ال ال في ال Packet وبيين ال IP بتاعها لو حصل match بينهم هترد ب ARP response Packet فيه ال MAC بتاعها وبالتالي يقدر ال Sender أليه يحدد ال ARP table الخاص به .

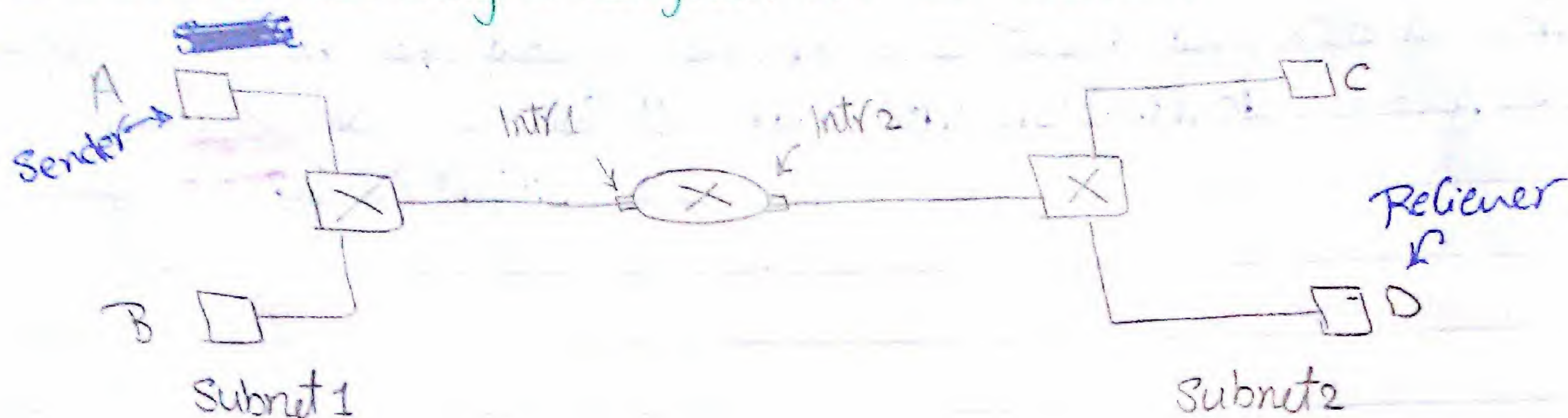
\* Notes: → The query ARP message is sent within a broadcast frame, whereas the response ARP message is sent within a standard frame.

- ~~ARP~~ ARP is Plug and play; ARP table gets built automatically without configuration by admin.



- If a host becomes disconnected from the subnet, its entry is eventually deleted from ARP tables.

- Scenario (Sending a datagram off the subnet)



\* ال Sender A هيجهز datagram ب dest. IP addr. و هيجهز encapsulation ب frame ب MAC addr ل intr 1 ل router ، و يمكن لبدا ال switch يجهز broadcast ل frame في Subnet 1 ، ل ال frame يوصل ل intr 1 هيجهز match ما بين ال MAC بتات والى موجود في ال frame وبالطاف ال adapter ل intr 1 هيجهز ال frame و يوصل ال IP ل datagram و هنا هيجهز دور ال forwarding table هيجهز ال forward ل datagram ( اللى ال IP ال D ) ل ال interface المناسب ل router 2 و هيجهز ال forward ل intr 2 و هنا يجهز ال encapsulation ب frame جديد ب ال MAC addr ل D و هيجهز broadcast في Subnet 2 .

\* Ethernet

→ Ethernet history

1. Ethernet with a hub-based star topology.
2. Switched Ethernet

→ Ethernet frame structure

Preamble	Dest. addr	Src. addr	Type	Data	CRC
----------	------------	-----------	------	------	-----



### ↳ Data field

- \* Carries IP datagram (46 → 1500 bytes)
- \* Maximum transmission unit (MTU) of Ethernet is 1500 bytes.
- لو ال IP datagram أقل من 46 byte ← لازم ال data field تنال  
لغاية ما توصل ل 46 والية دي اسمها **stuffing**

Fragmentation لو ال IP datagram أكبر من 1500 bytes هيجل

### ↳ Dest. addr.

- \* Contains dest. MAC address (6 bytes)

### ↳ Src. addr.

- \* Contains src. MAC address (6 bytes)

### ↳ Type

- \* indicates higher layer protocol type (2 byte)

### ↳ CRC

- \* allows the receiver to detect bit errors in the frame (4 byte)

### ↳ Preamble (8 byte)

- \* أول 7 bytes ليهم قيمة 10101010 وآخر byte قيمتها 11010101
- \* وظيفة أول 7 bytes أنهم يجلو **Synchronization** بين ال sender وال receiver

- \* آخر 2 <sup>bits</sup> byte الأخيرة (11) بتعمل indication ل receiver أن ال bytes ال بعد كده هتكون data

### - Notes:

- \* Ethernet isn't reliable<sup>(1)</sup>
- \* Ethernet technologies provide connectionless service to the network layer.<sup>(2)</sup>



لما (1) لا  
 Receiver B يستلم Frame من Sender A  
 CRC check على Frame كان فيه خطأ، هيجل discard  
 من غير ما يبعث لـ A أي acknowledgment وبالتالي الـ  
 Ethernet ← not reliable

لما (2) لا  
 Sender A علوز يبعث datagram لـ Receiver B  
 A هيجل encapsulation لـ datagram في Frame  
 وهيتبعه مع الـ LAN من غير أي handshaking في البداية  
 مع B.

\* Link layer switches

→ what's the role of the switch?

\* To receive incoming link layer frames and forward them onto outgoing links (forwarding)

— The switch is transparent to hosts and routers in the subnet.

فمثلاً الـ host هيجل addressing الـ Frame بالـ Receiver host addr.  
 وليس بعنوان الـ Switch intr.

— The rate at which frames arrive to any of the switches o/p interfaces may exceed the link capacity so switch o/p interfaces have buffers.

→ Forwarding and Filtering

\* Filtering → drop الـ Frame هيجل الـ drop

\* Forwarding → Frame هيجل الـ interfaces الـ هيجل الـ frame





\* Forwarding and Filtering are done with switch table

→ Switch table

Address	Interface	Time
MAC addr.	Switch intr. that leads to that MA addr.	time at which the entry was Placed in the table

→ Scenario (Forward & Filter in switch)

لفترض frame ل dest. addr معين وصل عند switch intr. ~~هنا فها تظهر 3 حالات :-~~

① أن مفيش entry في ال switch table ل addr. وهنا ال switch ~~هنا فها تظهر 3 حالات :-~~  
يعمل broadcast ل frame و كي ال interfaces كلها لا X.

② أن فيه entry ل addr. وال interface المقابل له هو ال X  
وهنا مفيش داعي ل frame ل نفس ال intr. ال الجاي  
منه فيحطه filtering و هيجل drop ل frame

③ أن فيه entry ل addr. وال interface المقابل له هو ال Y ( $Y \neq X$ )  
وهنا هيجل forwarding ل frame ل buffer ال المقابل  
ل Y intr.

→ How does the switch table get configured in the 1st place?

Switches are self-learning; its table's built automatically, dynamically without intervention from a network admin.

- This Process (Table configuration)'s accomplished as follows:-



- ① The table is empty.
- ② For each incoming frame received on an interface, the switch stores in its table (1) MAC addr. of the source, (2) the interface from which, the frame arrived, (3) The current time.
- ③ The switch deletes an address in the table if no frames are received with that address as src. addr.

— Switches are plug and play devices as they require no intervention from a network admin.

— Switches are full-duplex; any switch interface can send and receive at the same time.

### → Properties of link-layer switching:-

\* Elimination of collisions; the switches buffer frames and never transmit more than one frame on a segment at any one time.

\* Heterogeneous links as a switch isolates one link from another, different links on the LAN can operate at different speeds.

### \* Management

Ex: If an adapter continually sends Ethernet frames, a switch can detect the problem and disconnect the malfunctioning adapter.





## → Switches vs. Routers

### Switches

→ store and forward packets using MAC addr.

→ Layer-2 Packet switch

→ Link layer device

→ Plug and play devices

→ have forwarding table but differ in computation and configuration.

By self-learning

→ have high filtering and forwarding rates.

→ To prevent the cycling of broadcast frames, the active topology is restricted to a spanning tree.

### Routers

using network layer address

Layer-3 Packet switch

Network layer device

Need configurations

By Routing algorithms

have larger per packet processing time

Packets don't normally cycle through routers even when the network has redundant paths.

Packets are not restricted to a spanning tree and can use the best path between src. and dest.





## \* Virtual Local area networks (VLANs)

- Switch that supports VLANs allows multiple VLANs to be defined over a single physical local area network infrastructure.

(عازز أترق أكره VLAN مع نفس ال Switch)

- Slide 72 <sup>الرسمة</sup>

ال hosts مع نفس ال VLAN ~~موت~~ يتعاملوا كأنهم مع نفس ال Switch

## - Port-based VLAN

- \* Switch ports grouped so that single physical switch operates as multiple virtual switches using **switch management software** (table of Port to VLAN mappings)

## \* Supports traffic isolation:

ل بعض ال Frames من Ports مع ~~ال~~ VLAN ~~تقر~~ ، توصل ال Ports ال  
مع نفس ال VLAN فقط

ل عازز أترق أكره Communication ما بين VLANs ~~موت~~ فنتلقة :-

- \* By connecting a VLAN switch port to an external router and configure that port to belong to ~~these~~ all VLANs. (Vendors ease such configurations by building a single device that contains both a VLAN switch and a router)

## - ~~Scale~~ Scale up VLANs

\* Slide 74 <sup>الرسمة</sup>

\* نفس ال VLAN موجودة مع multiple physical switches

ل هاتج أكره Connection ما بين switches

## \* VLAN Trunking

Approach to ~~connect~~ interconnect VLAN switches

↳ A special port on each switch is configured as **trunk port** to interconnect ~~al~~ VLANs



↳ The trunk port belongs to all VLANs and frames sent to any VLAN're forwarded to the other switch over trunk port link.

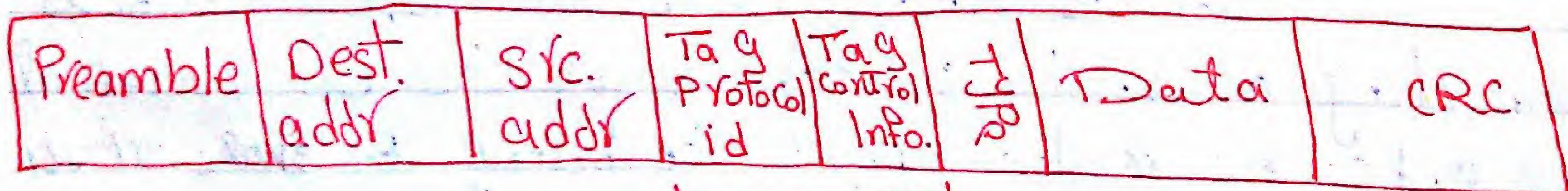
لذلك إذا كان الـ Switch يعرف الـ frame الذي هو موجود على الـ trunk port  
أنه يخص VLAN معينة ؟

\* IEEE's defined Ethernet Frame format 802.1Q for frames crossing a VLAN trunk.

### \* 802.1Q VLAN Frame Format

→ Consists of the standard Ethernet frame with 4 byte VLAN tag added into the header, carries the identity of the VLAN to which the frame belongs.

→ VLAN tag's added by the switch at the sending side of a VLAN trunk.



VLAN tag  
(4 bytes)

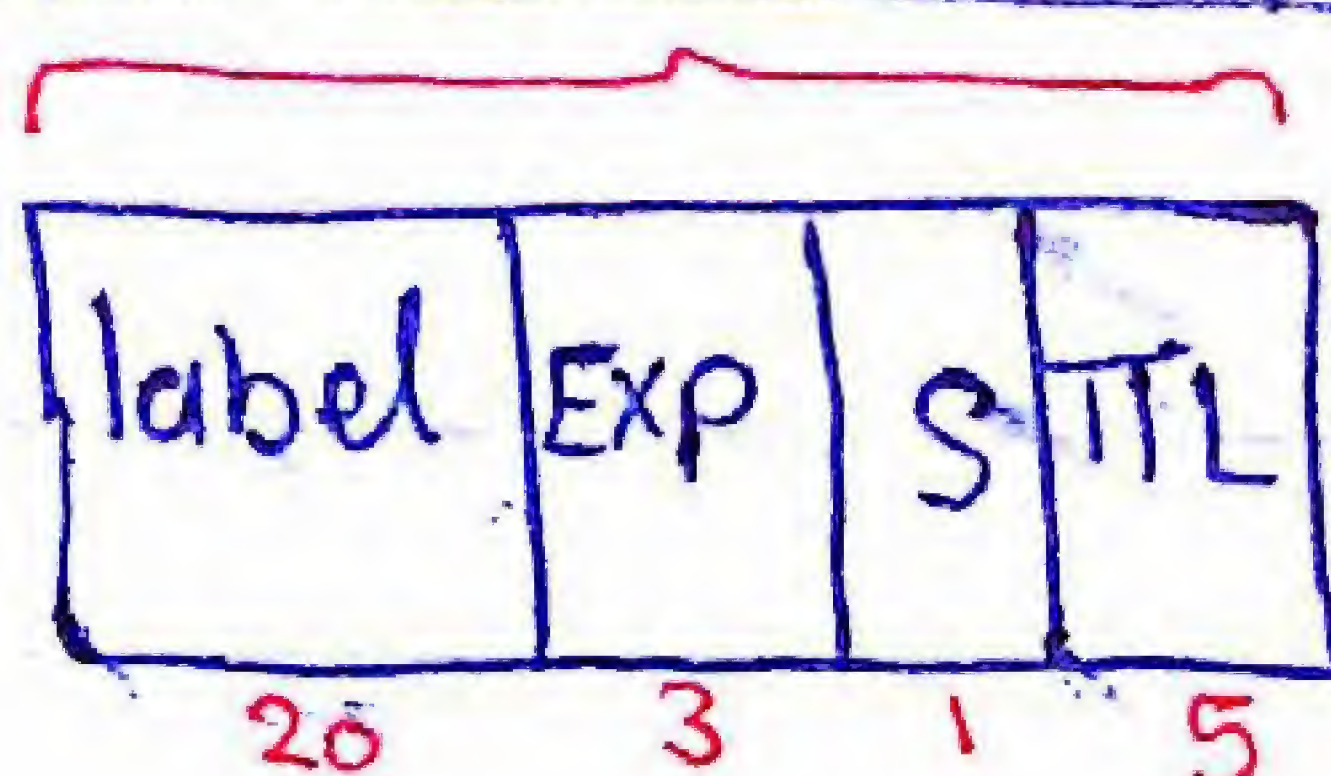
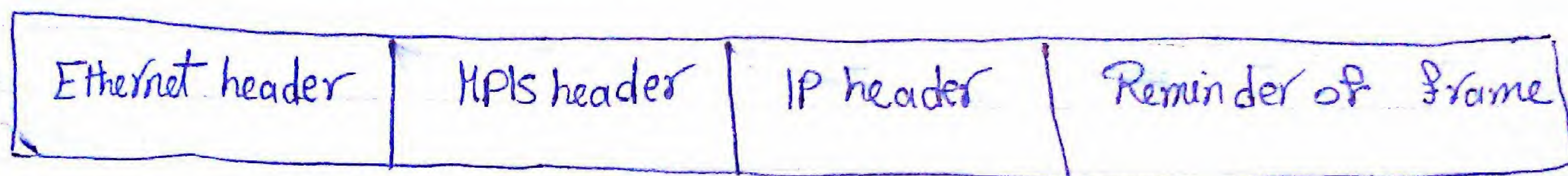


\*Link Virtualization : A network as a link layer  
 → Multiprotocol label switching (MPLS)

- MPLS evolved to improve the forwarding speed of IP routers by using a fixed-length label.

↳ allows routers to forward datagrams based on fixed-length label (rather than IP addr.) when possible. (don't inspect IP address)

- The format of a link layer frame that's handled by MPLS-capable router :-



EXP: experimental use

S: indicate the <sup>end</sup> of a series of MPLS headers

TTL: time-to-live

- MPLS forwarding table distinct from IP forwarding tables.

↳ MPLS forwarding table (for each router)

in label	out label	dest.	out interface

- MPLS-capable router is called also Label-Switched Router

\* How it works?

- Label Switched Router forwards an MPLS frame by looking up in its forwarding table and then immediately passing the datagram to the appropriate O/P interface, and need not to extract the IP and perform a look up of the longest Prefix match in the forwarding table.



## → MPLS vs. IP Paths

### \* IP Routing:

- Path to destination determined by destination address alone.

### \* MPLS Routing:

- Path to destination can be based on src and dest address.
- Fast reroute: Precompute routes in case of link failure.

## → slide 82

### \* Data Center networks

~~chall~~

- \* Thousands of hosts, often closely coupled.

#### \* Challenges:

- ↳ multiple apps, serve huge no. of clients.
- ↳ managing / Load balance
- ↳ Networking
- ↳ Rich Interconnection among switches and racks

#### \* Load balancer

- ↳ receives requests → direct workload to data center → return results to client  
(Hiding data center internal details)